

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

|  |    |   |
|--|----|---|
| (51) International Patent Classification 6 :<br><b>H04N 5/225</b>  | A1 | (11) International Publication Number: <b>WO 99/17543</b><br>(43) International Publication Date: 8 April 1999 (08.04.99)   |
| (21) International Application Number: <b>PCT/US98/13465</b><br>(22) International Filing Date: 25 June 1998 (25.06.98)<br><br>(30) Priority Data:<br>08/938,366 26 September 1997 (26.09.97) US   |    | (81) Designated States: AL, AM, AT, AT (Utility model), AU (Petty patent), AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, CZ (Utility model), DE, DE (Utility model), DK, DK (Utility model), EE, EE (Utility model), ES, FI, FI (Utility model), GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (Utility model), SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). |
| (71) Applicant: LIVE PICTURE, INC. [US/US]; Suite 300, 910 East Hamilton Avenue, Campbell, CA 95008 (US).<br>(72) Inventor: CHEN, Shenchang, Eric; 15951 Cerro Vista Court, Los Gatos, CA 95032 (US).<br>(74) Agents: SHEMWELL, Charles, E. et al.; Blakely, Sokoloff, Taylor & Zafman LLP, 7th floor, 12400 Wilshire Boulevard, Los Angeles, CA 90025 (US). |    | Published<br><i>With international search report.</i>   |
| (54) Title: VIRTUAL REALITY CAMERA   |    |   |
|  |    |   |
| (57) Abstract  |    |   |
| <p>A method and apparatus for creating and rendering multiple-view images. A camera (12) includes an image sensor (17) to receive images, sampling logic to digitize the images and a processor (19) programmed to combine the image based upon a spacial relationship between the images.</p>   |    |   |

BEST AVAILABLE COPY

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

|    |                          |    |                                       |    |   |    |                          |
|----|--------------------------|----|---------------------------------------|----|---|----|--------------------------|
| AL | Albania                  | ES | Spain                                 | LS | Lesotho                                   | SI | Slovenia                 |
| AM | Armenia                  | FI | Finland                               | LT | Lithuania                                 | SK | Slovakia                 |
| AT | Austria                  | FR | France                                | LU | Luxembourg                                | SN | Senegal                  |
| AU | Australia                | GA | Gabon                                 | LV | Latvia                                    | SZ | Swaziland                |
| AZ | Azerbaijan               | GB | United Kingdom                        | MC | Monaco                                    | TD | Chad                     |
| BA | Bosnia and Herzegovina   | GE | Georgia                               | MD | Republic of Moldova                       | TG | Togo                     |
| BB | Barbados                 | GH | Ghana                                 | MG | Madagascar                                | TJ | Tajikistan               |
| BE | Belgium                  | GN | Guinea                                | MK | The former Yugoslav Republic of Macedonia | TM | Turkmenistan             |
| BF | Burkina Faso             | GR | Greece                                | ML | Mali                                      | TR | Turkey                   |
| BG | Bulgaria                 | HU | Hungary                               | MN | Mongolia                                  | TT | Trinidad and Tobago      |
| BJ | Benin                    | IE | Ireland                               | MR | Mauritania                                | UA | Ukraine                  |
| BR | Brazil                   | IL | Israel                                | MW | Malawi                                    | UG | Uganda                   |
| BY | Belarus                  | IS | Iceland                               | MX | Mexico                                    | US | United States of America |
| CA | Canada                   | IT | Italy                                 | NE | Niger                                     | UZ | Uzbekistan               |
| CF | Central African Republic | JP | Japan                                 | NL | Netherlands                               | VN | Viet Nam                 |
| CG | Congo                    | KE | Kenya                                 | NO | Norway                                    | YU | Yugoslavia               |
| CH | Switzerland              | KG | Kyrgyzstan                            | NZ | New Zealand                               | ZW | Zimbabwe                 |
| CI | Côte d'Ivoire            | KP | Democratic People's Republic of Korea | PL | Poland                                    |    |                          |
| CM | Cameroon                 | KR | Republic of Korea                     | PT | Portugal                                  |    |                          |
| CN | China                    | KZ | Kazakhstan                            | RO | Romania                                   |    |                          |
| CU | Cuba                     | LC | Saint Lucia                           | RU | Russian Federation                        |    |                          |
| CZ | Czech Republic           | LI | Liechtenstein                         | SD | Sudan                                     |    |                          |
| DE | Germany                  | LK | Sri Lanka                             | SE | Sweden                                    |    |                          |
| DK | Denmark                  | LR | Liberia                               | SG | Singapore                                 |    |                          |

## VIRTUAL REALITY CAMERA

### FIELD OF THE INVENTION

The present invention relates to the field of photography, and more particularly to a camera that combines images based on a spatial relationship between the images.

### BACKGROUND OF THE INVENTION

A panoramic image of a scene has traditionally been created by rotating a vertical slit camera about an optical center. Using this technique, film at the optical center is continuously exposed to create a wide field of view (e.g., a 360° field of view). Because of their specialized design, however, vertical slit cameras are relatively expensive. Further, because the panoramic image is captured in a continuous rotation of the camera, it is difficult to adjust the camera to account for changes in the scene, such as lighting or focal depth, as the camera is rotated.

In a more modern technique for creating panoramic images, called "image stitching", a scene is photographed from different camera orientations to obtain a set of discrete images. The discrete images of the scene are then transferred to a computer which executes application software to blend the discrete images into a panoramic image.

After the panoramic image is created, application software may be executed to render user-specified portions of the panoramic image onto a

display. The effect is to create a virtual environment that can be navigated by a user. Using a mouse, keyboard, headset or other input device, the user can pan about the virtual environment and zoom in or out to view objects of interest.

One disadvantage of existing image stitching techniques is that photographed images must be transferred from the camera to the computer before they can be stitched together to create a navigable panoramic image. For example, with a conventional exposed-film camera, film must be exposed, developed, printed and digitized (e.g., using a digital scanner) to obtain a set of images that can be stitched into a panoramic image. In a digital camera, the process is less cumbersome, but images must still be transferred to a computer to be stitched into a panoramic view.

Another disadvantage of existing image stitching techniques is that the orientation of the camera used to photograph each discrete image is typically unknown. This makes it more difficult to stitch the discrete images into a panoramic image because the spatial relationship between the constituent images of the panoramic image are determined, at least partly, based on the respective orientations of the camera at which they were captured. In order to determine the spatial relationship between a set of images that are to be stitched into a panoramic image, application software must be executed to prompt the user for assistance, hunt for common features in the images, or both.

Yet another disadvantage of existing image stitching techniques is that it is usually not possible to determine whether there are missing

views in the set of images used to create the panoramic image until after the images have been transferred to the computer and stitched. Depending on the subject of the panoramic image, it may be inconvenient or impossible to recreate the scene necessary to obtain the missing view. Because of the difficulty determining whether a complete set of images has been captured, images to be combined into a panoramic image are typically photographed with conservative overlap to avoid gaps in the panoramic image. Because there is more redundancy in the captured images, however, a greater number of images must be obtained to produce the panoramic view. For conventional film cameras, this means that more film must be exposed, developed, printed and scanned to produce a panoramic image than if less conservative image overlap were possible. For digital cameras, more memory must typically be provided to hold the larger number of images that must be captured than if less conservative image overlap were possible.

#### SUMMARY OF THE INVENTION

A method and apparatus for creating and rendering multiple-view images are disclosed. Images are received on the image sensor of a camera and digitized by sampling logic in the camera. The digitized images are combined by a programmed processor in the camera based upon a spatial relationship between the images.

#### BRIEF DESCRIPTION OF THE DRAWINGS

The present invention is illustrated by way of example and not

limitation in the figures of the accompanying drawings in which like references indicate similar elements and in which:

Fig. 1 is a block diagram of a virtual reality (VR) camera.

Fig. 2 illustrates the use of a VR camera to generate a panoramic image.

Fig. 3 illustrates the use of a VR camera to generate a composite image of a surface.

Fig. 4 illustrates the use of a VR camera to generate an object image.

Fig. 5 illustrates control inputs on a VR camera according.

Fig. 6 illustrates the use of a VR camera to overlay a video feed over a previously recorded scene.

Fig. 7 is a block diagram of a stereo VR camera.

Fig. 8 is a diagram of a method according to one embodiment of the present invention.

Fig. 9 is a diagram of a method according to an alternate embodiment of the present invention.

#### DETAILED DESCRIPTION

According to the present invention, a virtual reality (VR) camera is provided to create and render panoramic images and other multiple-view images. In one embodiment, the VR camera includes a sensor to detect the camera orientation at which images in a scene are captured. A computer within the VR camera combines the images of the scene into a panoramic image based, at least partly, on the respective camera orientations at which the images were captured. A display in the VR

camera is used to view the panoramic image. In one embodiment of the present invention, the orientation of the VR camera is used to select which portion of the panoramic image is displayed so that a user can effectively pan about the panoramic image by changing the orientation of the camera.

Fig. 1 is a block diagram of a VR camera 12 according to one embodiment of the present invention. VR camera 12 may be either a video camera or a still-image camera and includes an optic 15, an image acquisition unit (IAU) 17, an orientation/position sensor (O/P sensor) 21, one or more user input panels 23, a processor 19, a non-volatile program code storage 24, a memory 25, a non-volatile data storage 26 and a display 27.

The optic 15 generally includes an automatically or manually focused lens and an aperture having a diameter that is adjustable to allow more or less light to pass. The lens projects a focused image through the aperture and onto an image sensor in the IAU 17. The image sensor is typically a charge-coupled device (CCD) that is sampled by sampling logic in the IAU 17 to develop a digitized version of the image. The digitized image may then be read directly by the processor 19 or transferred from the IAU 17 to the memory 25 for later access by the processor 19. Although a CCD sensor has been described, any type of image sensor that can be sampled to generate digitized images may be used without departing from the scope of the present invention.

In one embodiment of the present invention, the processor 19 fetches and executes program code stored in the code storage 24 to

implement a logic unit capable of obtaining the image from the IAU 17 (which may include sampling the image sensor), receiving orientation and position information from the O/P sensor 21, receiving input from the one or more user input panels 23 and outputting image data to the display 27. It will be appreciated that multiple processors, or hard-wired logic may alternatively be used to perform these functions. The memory 25 is provided for temporary storage of program variables and image data, and the non-volatile image storage 26 is provided for more permanent storage of image data. The non-volatile storage 26 may include a removable storage element, such as a magnetic disk or tape, to allow panoramic and other multiple-view images created using the VR camera 12 to be stored indefinitely.

The O/P sensor 21 is used to detect the orientation and position of the VR camera 12. The orientation of the VR camera 12 (i.e., pitch, yaw and roll) may be determined relative to an arbitrary starting orientation or relative to a fixed reference (e.g., earth's gravitational and magnetic fields). For example, an electronic level of the type commonly used in virtual reality headsets can be used to detect camera pitch and roll (rotation about horizontal axes), and an electronic compass can be used to detect camera yaw (rotation about a vertical axis). As discussed below, by recording the orientation of the VR camera 12 at which each of a set of discrete images is captured, the VR camera 12 can automatically determine the spatial relationship between the discrete images and combine the images into a panoramic image, planar composite image, object image or any other type of multiple-view image.

Still referring to Fig. 1, when a panoramic image (or other multiple-view image) is displayed on display 27, changes in camera orientation are detected via the O/P sensor 21 and interpreted by the processor 19 as requests to pan about the panoramic image. Thus, by rotating the VR camera 12 in different directions, a user can view different portions of the previously generated panoramic image on the display 27. The VR camera's display 27 becomes, in effect, a window into a virtual environment that has been created in the VR camera 12.

In one embodiment of the present invention, the position of the VR camera 12 in a three-dimensional (3D) space is determined relative to an arbitrary or absolute reference. This is accomplished, for example, by including in the O/P sensor 21 accelerometers or other devices to detect translation of VR the camera 12 relative to an arbitrary starting point. As another example, the absolute position of the VR camera 12 may be determined including in the O/P sensor 21 a sensor that communicates with a global positioning system (GPS). GPS is well known to those of ordinary skill in the positioning and tracking arts. As discussed below, the ability to detect translation of the VR camera 12 between image capture positions is useful for combining discrete images to produce a composite image of a surface.

It will be appreciated from the foregoing discussion that the O/P sensor 21 need not include both an orientation sensor and a position sensor, depending on the application of the VR camera 12. For example, to create and render a panoramic image, it is usually necessary to change the angular orientation of the VR camera 12 only. Consequently, in one

embodiment of the present invention, the O/P sensor 21 is an orientation sensor only. Other combinations of sensors may be used without departing from the scope of the present invention.

Still referring to Fig. 1, the one or more user input panels 23 may be used to provide user control over such conventional camera functions as focus and zoom (and, at least in the case of a still camera, aperture size, shutter speed, etc.). As discussed below, the input panels 23 may also be used to receive user requests to pan about or zoom in and out on a panoramic image or other multiple-view image. Further, the input panels 23 may be used to receive user requests to set certain image capture parameters, including parameters that indicate the type of composite image to be produced, whether certain features are enabled, and so forth. It will be appreciated that focus and other camera settings may be adjusted using a traditional lens dial instead of an input panel 23. Similarly, other types of user input devices and techniques, including, but not limited to, user rotation and translation of the VR camera 12, may be used to receive requests to pan about or zoom in or out on an image.

The display 27 is typically a liquid crystal display (LCD) but may be any type of display that can be included in the VR camera 12, including a cathode-ray tube display. Further, as discussed below, the display 27 may be a stereo display designed to present left and right stereo images to the left and right eyes, respectively, of the user.

Fig. 2 illustrates use of the VR camera 12 of Fig. 1 to generate a panoramic image 41. A panoramic image is an image that represents a wide-angle view of a scene and is one of a class of images referred to

herein as multiple-view images. A multiple-view image is an image or collection of images that is displayed in user-selected portions.

To create panoramic image 41, a set of discrete images 35 is first obtained by capturing images of an environment 31 at different camera orientations. With a still camera, capturing images means taking photographs. With a video camera, capturing image refers to generating one or more video frames of each of the discrete images.

For ease of understanding, the environment 31 is depicted in Fig. 2 as being an enclosed space but this is not necessary. In order to avoid gaps in the panoramic image, the camera is oriented such that each captured image overlaps the preceding captured image. This is indicated by the overlapped regions 33. The orientation of the VR camera is detected via the O/P sensor (e.g., element 21 of Fig. 1) and recorded for each of the discrete images 35.

In one still-image camera embodiment of the present invention, as the user pans the camera about the environment 31, the orientation sensor is monitored by the processor (e.g., element 19 of Fig. 1) to determine when the next photograph should be snapped. That is, the VR camera assists the photographer in determining the camera orientation at which each new discrete image 35 is to be snapped by signaling the photographer (e.g., by turning on a beeper or a light) when region of overlap 33 is within a target size. Note that the VR camera may be programmed to determine when the region of overlap 33 is within a target size not only for camera yaw, but also for camera pitch or roll. In another embodiment of the present invention, the VR camera may be user-

configured (e.g., via a control panel 23 input) to automatically snap a photograph whenever it detects sufficient change in orientation. In both manual and automatic image acquisition modes, the difference between camera orientations at which successive photographs are acquired may be input by the user or automatically determined by the VR camera based upon the camera's angle of view and the distance between the camera and subject.

In a video camera embodiment of the present invention, the orientation sensor may be used to control the rate at which video frames are generated so that frames are generated only when the O/P sensor indicates sufficient change in orientation (much like the automatic image acquisition mode of the still camera discussed above), or video frames may be generated at standard rates with redundant frames being combined or discarded during the stitching process.

As stated above, the overlapping discrete images 35 can be combined based on their spatial relationship to form a panoramic image 41.

Although the discrete images 35 are shown as being a single row of images (indicating that the images were all captured at approximately same pitch angle), additional rows of images at higher or lower pitch angles could also have been obtained. Further, because the VR camera will typically be hand held (although a tripod may be used), a certain amount of angular error is incurred when the scene is recorded. This angular error is indicated in Fig. 2 by the slightly different pitch and roll orientation of the discrete images 35 relative to one another, and must be accounted for when the images are combined to form the panoramic image 41.

After the discrete images 35 have been captured and stored in the memory of the camera (or at least two of the discrete image have been captured and stored), program code is executed in the VR camera to combine the discrete images 35 into the panoramic image 41. This is accomplished by determining a spatial relationship between the discrete images 35 based on the camera orientation information recorded for each image 35, or based on common features in the overlapping regions of the images 35, or based on a combination of the two techniques.

One technique for determining a spatial relationship between images based on common features in the images is to "cross-correlate" the images. Consider, for example, two images having an unknown translational offset relative to one another. The images can be cross-correlated by "sliding" one image over the other image one step (e.g., one pixel) at a time and generating a cross-correlation value at each sliding step. Each cross-correlation value is generated by performing a combination of arithmetic operations on the pixel values within the overlapping regions of the two images. The offset that corresponds to the sliding step providing the highest correlation value is found to be the offset of the two images. Cross-correlation can be applied to finding offsets in more than one direction or to determine other unknown transformational parameters, such as rotation or scaling. Techniques other than cross-correlation, such as pattern matching, can also be used to find unknown image offsets and other transformational parameters.

Based on the spatial relationship between the discrete images 35, the images 35 are mapped onto respective regions of a smooth surface such as

a sphere or cylinder. The regions of overlap 33 are blended in the surface mapping. Depending on the geometry of the surface used, pixels in the discrete images 35 must be repositioned relative to one another in order to produce a two-dimensional pixel-map of the panoramic image 41. For example, if the discrete images 35 are mapped onto a cylinder 37 to produce the panoramic image 41, then horizontal lines in the discrete images 35 will become curved when mapped onto the cylinder 37 with the degree of curvature being determined by latitude of the horizontal lines above the cylindrical equator. Thus, stitching the discrete images 35 together to generate a panoramic image 41 typically involves mathematical transformation of pixels to produce a panoramic image 41 that can be rendered without distortion.

Fig. 3 illustrates the use of the VR camera 12 to generate a composite image of a surface 55 that is too detailed to be adequately represented in a single photograph. Examples of such surfaces include a white-board having notes on it, a painting, an inscribed monument (e.g., the Viet Nam War Memorial), and so forth.

As indicated in Fig. 3, multiple discrete images 57 of the surface 55 are obtained by translating the VR camera 12 between a series of positions and capturing a portion of the surface 55 at each position. According to one embodiment of the present invention, the position of the VR camera 12 is obtained from the position sensing portion of the O/P sensor (element 21 of Fig. 1) and recorded for each discrete image 57. This allows the spatial relationship between the discrete images 57 to be determined no matter the order in which the images 57 are obtained. Consequently,

the VR camera is able to generate an accurate composite image 59 of the complete surface 55 regardless of the order in which the discrete images 57 are captured. In the case of a still image camera, the position sensor can be used to signal the user when the VR camera 12 has been sufficiently translated to take a new photograph. Alternatively, the VR camera may be user-configured to automatically snap photographs as the VR camera 12 is swept across the surface 55. In the case of a video camera, the position sensor can be used to control when each new video frame is generated, or video frames may be generated at the standard rate and then blended or discarded based on position information associated with each.

After two or more of the discrete images 57 have been stored in the memory of the VR camera 12, program code can be executed to combine the images into a composite image 59 based on the position information recorded for each discrete image 57, or based on common features in overlapping regions of the discrete images 57, or both. After the discrete images 57 have been combined into a composite image 59, the user may view different portions of the composite image 59 on the VR camera's display by changing the orientation of the VR camera 12 or by using controls on a user input panel. By zooming in at a selected portion of the image, text on a white-board, artwork detail, inscriptions on a monument, etc. may be easily viewed. Thus, the VR camera 12 provides a simple and powerful way to digitize and render high resolution surfaces with a lower resolution camera. Composite images of such surfaces are referred to herein as "planar composite images", to distinguish them from panoramic images.

Fig. 4 illustrates yet another application of the VR camera. In this case the VR camera is used to combine images into an object image 67. An object image is a set of discrete images that are spatially related to one another, but which have not been stitched together to form a composite image. The combination of images into an object image is accomplished by providing information indicating the location of the discrete images relative to one another and not by creating a separate composite image.

As shown in Fig. 4, images of an object 61 are captured from surrounding points of view 63. Though not shown in the plan view of the object 61, the VR camera may also be moved over or under the object 61, or may be raised or tilted to capture images of the object 61 at different heights. For example, the first floor of a multiple-story building could be captured in one sequence of video frames (or photographs), the second floor in a second sequence of video frames, and so forth. If the VR camera is maintained at an approximately fixed distance from the object 61, the orientation of the VR camera alone may be recorded to establish the spatial relationship between the discrete images 65. If the object is filmed (or photographed) from positions that are not equidistant to the object 61, it may be necessary to record both the position and orientation of the VR camera for each discrete image 65 in order to produce a coherent object image 67.

After two or more discrete images 65 of object 61 have been obtained, they can be combined based upon the spatial relationship between them to form an object image 67. As stated above, combining the discrete images 65 to form an object image 67 typically does not involve

stitching the discrete images 65 and is instead accomplished by associating with each of the discrete images 65 information that indicates the image's spatial location in the object image 67 relative to other images in the object image 67. This can be accomplished, for example, by generating a data structure having one member for each discrete image 65 and which indicates neighboring images and their angular or positional proximity. Once the object image 67 is created, the user can pan through the images 65 by changing the orientation of the camera. Incremental changes in orientation can be used to select an image in the object image 67 that neighbors a previously displayed image. To the user, rendering of the object image 67 in this manner provides a sense of moving around, over and under the object of interest.

According to another embodiment of the present invention, the relative spatial location of each image in the object image 67 an object image is provided by creating a data structure containing the camera orientation information recorded for each discrete image 65. To select a particular image in the object image 67, the user orients the VR camera in the direction that was used to capture the image. The VR camera's processor detects the orientation via the orientation sensor, and then searches the data structure to identify the discrete image 65 having a recorded orientation most nearly matching the input orientation. The identified image 65 is then displayed on the VR camera's display.

Fig. 5 depicts a VR camera 12 that is equipped with a number of control buttons that are included in user input panels 23a and 23b. The buttons provided in user-input panel 23a vary depending on whether VR

camera 12 is a video camera or a still-image camera. For example, in a still-image camera, panel 23a may include shutter speed and aperture control buttons, among others, to manage the quality of the photographed image. In a video camera, user input panel 23a may include, for example, zoom and focus control. User input panel 23a may also include mode control buttons to allow a user to select certain modes and options associated with creating and rendering virtual reality images. In one embodiment, for example, mode control buttons may be used to select a panoramic image capture mode, planar composite image capture mode or object image capture mode. Generally, any feature of the VR camera that can be selected, enabled or disabled may be controlled using the mode control buttons.

According to one embodiment of the present invention, view control buttons Right /Left, Up/Down and Zoom are provided in user input panel 23b to allow the user to select which portion of a panoramic image, planar composite image, object image or other multiple-view image is presented on display 27. When the user presses the Right button, for example, view control logic in the camera detects the input and causes the displayed view of a composite image or object image to pan right. When the user presses the Zoom+ button, the view control logic causes the displayed image to be magnified. The view control logic may be implemented by a programmed processor (e.g., element 19 of Fig. 1), or by dedicated hardware. In one embodiment of the present invention, the view control logic will respond either to user input via panel 23b or to changes in camera orientation. Alternatively, the camera may be

configured such that in one mode, view control is achieved by changing the VR camera orientation, and in another mode, view control is achieved via the user input panel 23b. In both cases, the user is provided with alternate ways to select a view of a multiple-view image.

Fig. 6 illustrates yet another application of the VR camera 12 of the present invention. In this application, a video signal captured via the IAU (element 17 of Fig. 1)a is superimposed on a previously recorded scene using a chroma-key color replacement technique. For example, an individual 83 standing in front of a blue background 82 may be recorded using the VR camera 12 to generate a live video signal. Program code in the VR camera 12 may then be executed to implement an overlay function that replaces pixels in a displayed scene with non-blue pixels from the live video. The effect is to place the subject 83 of the live video in the previously generated scene. According to one embodiment of the present invention, the user may pan about a panoramic image on display 27 to locate a portion of the image into which the live video is to be inserted, then snap the overlaid subject of the video image into the scene. In effect, the later received image is made part of the earlier recorded panoramic image (or other multiple-view image) and the combined images can be permanently stored as a single recorded video or still image.

Fig. 7 is a block diagram of a VR camera 112 that is used to receive and process stereo images. As shown, the optic 115 includes both left and right channels (108, 107) for receiving respective left and right images. Typically the left and right images are of the same subject but from spatially differentiated viewpoints. This way a 3D view of the subject is

captured. According to one embodiment of the present invention, the left and right images 108 and 107 are projected onto opposing halves of an image sensor in the IAU 117 where they are sampled by the processor 19 and stored in memory 25. Alternatively, multiple image sensors and associated sampling circuitry may be provided in the IAU 117. In either case, the left and right images are associated with orientation/position information obtained from the O/P sensor 21 in the manner described above, and stored in the memory 25. After two or more discrete images have been obtained, the processor may execute program code in the non-volatile code storage 24 to combine the left images into a left composite image and the right images into a right composite image. In an object image application, the processor combines the right and left images into respective right and left object images.

As shown in Fig. 7, a stereo display 127 is provided to allow a 3D view of a scene to be displayed. For example, a polarized LCD display that relies on the different viewing angles of the left and right eyes of an observer may be used. The different viewing angles of the observer's left and right eyes causes different images to be perceived by the left and right eyes. Consequently, based on an orientation/position of the camera, or a view select input from the user, a selected portion of the left composite image (or object image) is presented to the left eye and a selected portion of the right composite image (or object image) is presented to the right eye.

As with the VR camera 12 described above, live stereo video received in the IAU 117 of the stereo VR camera 112 may be overlaid on a previously generated composite image or object image. The left and right

video components of the live stereo video may be superimposed over the left and right composite or object images, respectively. Consequently, the user may view live video subjects in 3D as though they were present in the previously recorded 3D scene. A stereo photograph may also be overlaid on an earlier recorded composite image or object image.

Fig. 8 is a diagram of a method according to one embodiment of the present invention. At step 141, a set of discrete images are received in the camera. The images are digitized at step 143. Based upon a spatial relationship between the digitized images, the digitized images are combined to produce a multiple-view image at step 145. Then, at step 145, at least a portion of the multiple-view image is displayed on a display of the camera.

It will be appreciated from the foregoing description of the present invention that the steps of receiving (141), digitizing (143) and combining (145) may be performed on an image by image basis so that each image is received, digitized and combined with one or more previously received and digitized images before a next image is received and digitized.

A method of generating of a multiple-view image on a discrete image by discrete image basis shown in Fig. 9. At step 151, a discrete image<sub>i</sub> is received, where i ranges from 0 to N. At step 153, image<sub>i</sub> is digitized, and i is incremented at step 157. If i is determined to be less than or equal to one at step 159, execution loops back to step 151 to receive the next discrete image<sub>i</sub>. If i is greater than one, then at step 161 digitized image<sub>i</sub> is combined with one or more previously digitized images based on a spatial relationship between the digitized image<sub>i</sub> and the one or more

previously digitized images to produce a multiple-view image. If it is determined that a final image has been received and digitized, (arbitrarily shown as N in step 163) the method is exited. It will be appreciated that the determination as to whether a final image has been received may be made in a number of ways, including: detecting that a predetermined number of images have been received, digitized and combined; or receiving a signal from the user or an internally generated signal indicating that a desired or threshold number of images have been received, digitized and combined into the multiple-view image. Also, according to one embodiment of the present invention, the user may select a portion of the multiple-view image for viewing any time after an initial combining step 159 has been performed.

In the foregoing specification, the invention has been described with reference to specific exemplary embodiments thereof. It will, however, be evident that various modifications and changes may be made thereto without departing from the broader spirit and scope of the invention as set forth in the appended claims. The specification and drawings are, accordingly to be regarded in an illustrative rather than a restrictive sense.

**CLAIMS****What is claimed is:**

1. A camera comprising:  
an image sensor to receive images;  
sampling logic to digitize the images; and  
a processor programmed to combine the digitized images based upon  
a spatial relationship between the digitized images.
2. The camera of claim 1 wherein the spatial relationship between the  
digitized images is determined based on one or more features in the  
digitized images.
3. The camera of claim 1 further comprising an orientation sensor to  
detect respective orientations of the camera at which the images are  
received and wherein the spatial relationship between the digitized  
images is based on the orientations of the camera.
4. The camera of claim 1 further comprising a positional sensor to detect  
respective positions of the camera at which the images are received  
and wherein the spatial relationship between the digitized images is  
based on the positions of the camera.

5. The camera of claim 1 wherein the processor combines the digitized images into a composite image.
6. The camera of claim 5 further comprising:  
a display; and  
view control logic to select at least a portion of the composite image to be displayed on the display.
7. The camera of claim 6 further comprising a user interface to receive from a camera user a request to display the portion of the composite image and wherein the view control logic selects the portion of the composite image to be displayed based on the request, the view control logic being implemented by the processor.
8. The camera of claim 6 further comprising an orientation sensor to detect an orientation of the camera and wherein the view control logic selects the portion of the composite image to be displayed based upon the orientation of the camera.
9. The camera of claim 6 further comprising a position sensor to detect a position of the camera and wherein the view control logic selects the portion of the composite image to be displayed based upon the position of the camera.

10. The camera of claim 6 further comprising overlay logic to replace a portion of the composite image with a portion of an additional image received on the image sensor.
11. The camera of claim 1 wherein the processor is programmed to combine the images into an object image.
12. The camera of claim 11 wherein the images are combined into an object image by associating with each of the images information indicating respective orientation of the camera at which the image is captured.
13. The camera of claim 11 further comprising:  
an orientation sensor to detect an orientation of the camera;  
view control logic to select based on the orientation of the camera a portion of the object image; and  
a display to display the selected portion of the object image.
14. The camera of claim 1 wherein the images received by the image sensor are stereo image pairs with each stereo image pair including a left image and a right image, and wherein the processor is programmed to combine left images in the stereo image pairs into a left composite image and right images in the stereo image pairs into a right composite image.

15. The camera of claim 14 further comprising an optical assembly including left and right optical channels to project the left and right images, respectively, onto the image sensor, the left and right optical channels being disposed to provide spatially differentiated views of a subject.
16. The camera of claim 14 further comprising:  
a stereo display to present left and right images in the stereo image pairs to left and right eyes, respectively, of an observer; and  
view control logic to select at least respective portions of the left and right composite images to be displayed on the stereo display.
17. The camera of claim 1 wherein the camera is a video camera and wherein the sampling logic digitizes the images at a predetermined rate.
18. A camera comprising:  
means for receiving images;  
means for digitizing the images; and  
means for combining the digitized images based upon a spatial relationship between the digitized images.
19. The camera of claim 18 further comprising means for detecting respective orientations of the camera at which the images are

received and wherein the spatial relationship is based on the orientations of the camera.

20. The camera of claim 18 further comprising means for detecting respective positions of the camera at which the images are received and wherein the spatial relationship is based on the positions of the camera.
21. The camera of claim 18 wherein the means for combining combines the images into a composite image and wherein the camera further comprises:  
means for displaying; and  
means for selecting at least a portion of the composite image to be displayed on the means for displaying.
22. The camera of claim 21 further comprising means for detecting an orientation of the camera and wherein the means for selecting selects the portion of the composite image to be displayed based upon the orientation of the camera.
23. The camera of claim 21 further comprising means for detecting a position of the camera and wherein the means for selecting selects the portion of the composite image to be displayed based upon the position of the camera.

24. The camera of claim 21 further comprising means for replacing a portion of the composite image with a portion of an additional image received on the means for receiving images.
25. The camera of claim 18 wherein the means for combining combines the images into an object image.
26. The camera of claim 25 wherein the means for combining combines the images into an object image by associating with each of the images information indicating a respective orientation of the camera at which the image is captured.
27. The camera of claim 25 further comprising:  
means for detecting an orientation of the camera;  
means for selecting based on the orientation of the camera a portion of the object image; and  
means for displaying the selected portion of the object image.
28. A method comprising the steps of:  
receiving images in a camera;  
digitizing the images;  
combining the digitized images based upon a spatial relationship between the digitized images to produce a composite image; and

displaying at least a portion of the composite image on a display of the camera.

29. The method of claim 28 further comprising the steps of:  
detecting respective orientations of the camera at which the images are received; and  
determining the spatial relationship between the digitized images based on the respective orientations.
30. The method of claim 28 further comprising the steps of:  
detecting respective positions of the camera at which the images are received; and  
determining the spatial relationship between the digitized images based on the respective positions.
31. The method of claim 28 further comprising the step of determining the spatial relationship between the digitized images based on one or more features in the digitized images.
32. The method of claim 28 further comprising the steps of:  
detecting an orientation of the camera; and  
selecting the portion of the composite image to be displayed based upon the orientation of the camera.

33. The method of claim 28 further comprising the steps of:  
detecting a position of the camera; and  
selecting the portion of the composite image to be displayed based  
upon the position of the camera.
34. The method of claim 28 further comprising the steps of:  
receiving an additional image;  
digitizing the additional image; and  
replacing a portion of the composite image with a portion of the  
additional image.
35. The method of claim 28 wherein the step of receiving images  
includes the step of receiving stereo image pairs with each stereo  
image pair including a left image and a right image, and wherein the  
step of combining includes the step of combining left images in the  
stereo image pairs into a left composite image and right images in the  
stereo image pairs into a right composite image.
36. The method of claim 35 wherein the step of displaying at least a  
portion of the composite image includes the steps of:  
displaying at least a portion of the left composite image to a left eye of  
an observer; and  
displaying at least a portion of the right composite image to a right

eye of an observer.

37. A method comprising the steps of:  
receiving images in a camera;  
digitizing the images; and  
combining the digitized images based upon a spatial relationship  
between the digitized images to produce an object image.
38. The camera of claim 37 wherein the step of combining the digitized  
images to produce an object image includes the step of associating  
with each of the digitized images information indicating respective  
orientation of the camera at which the digitized image is captured.
39. The method of claim 37 further comprising the steps of:  
detecting an orientation of the camera;  
selecting based on the orientation of the camera a portion of the  
object image; and  
displaying the selected portion of the object image on a display of the  
camera.
40. A camera comprising:  
an image sensor to receive images; and

a sensor to detect respective orientations of the camera at which the images are received.

1/9

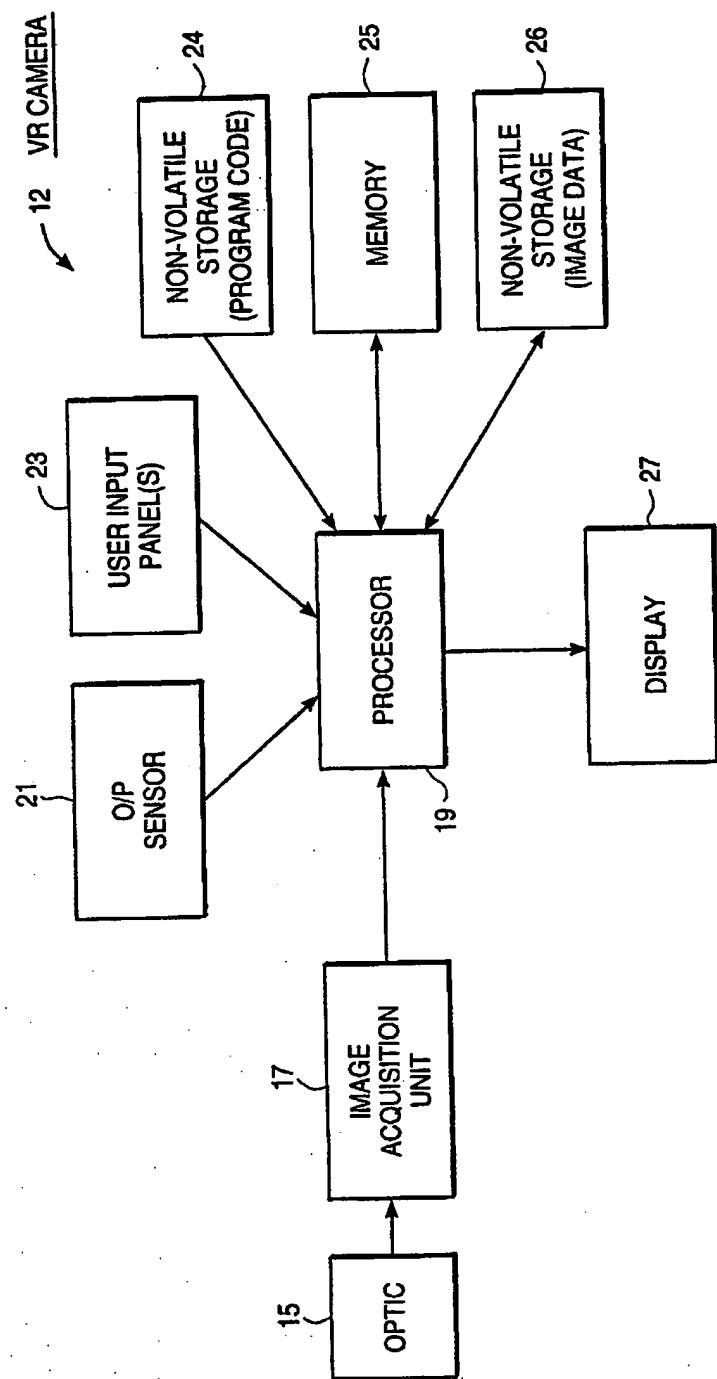


FIG. 1

SUBSTITUTE SHEET (RULE 26)

2/9

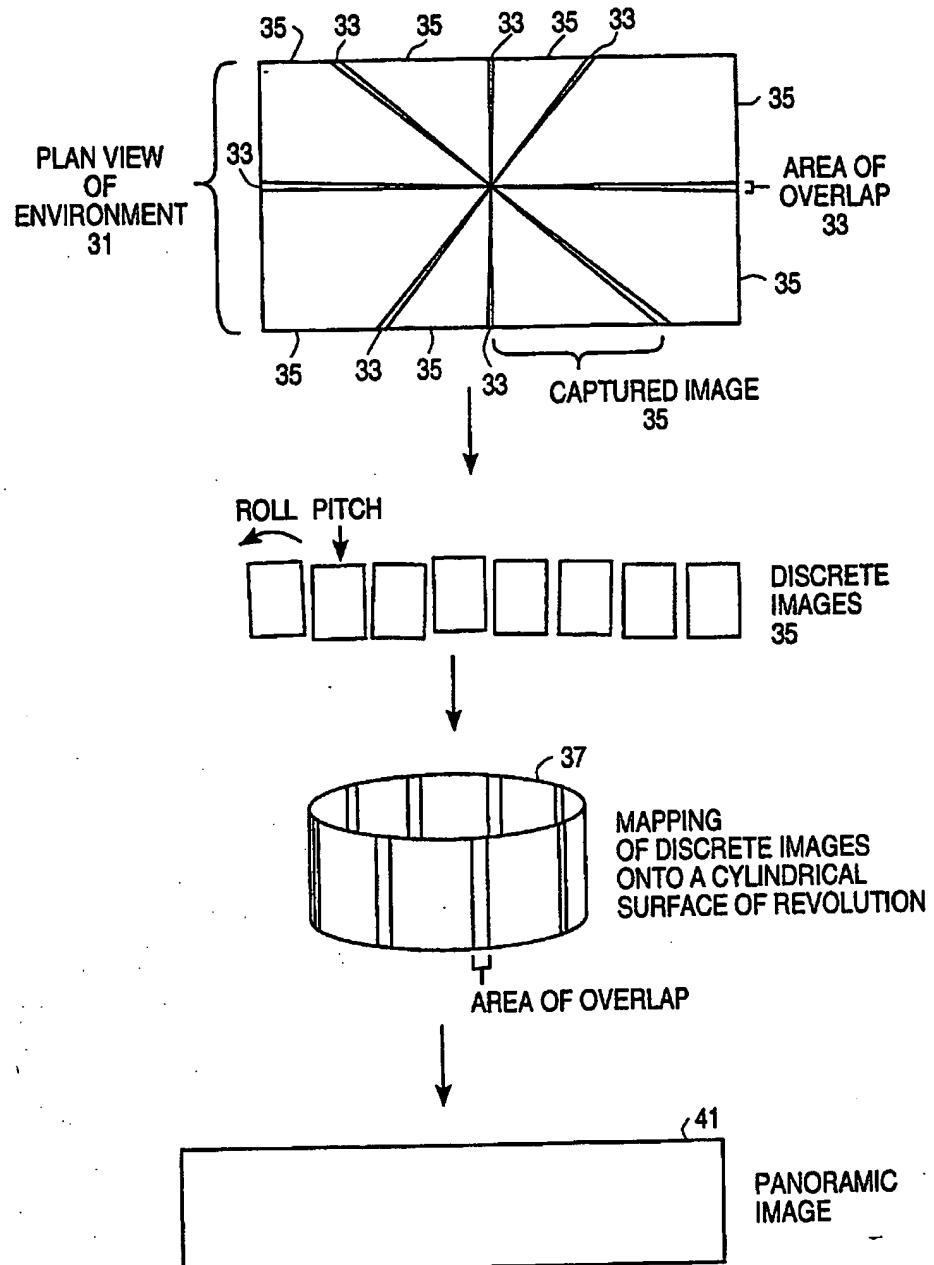


FIG. 2  
SUBSTITUTE SHEET (RULE 26)

3/9

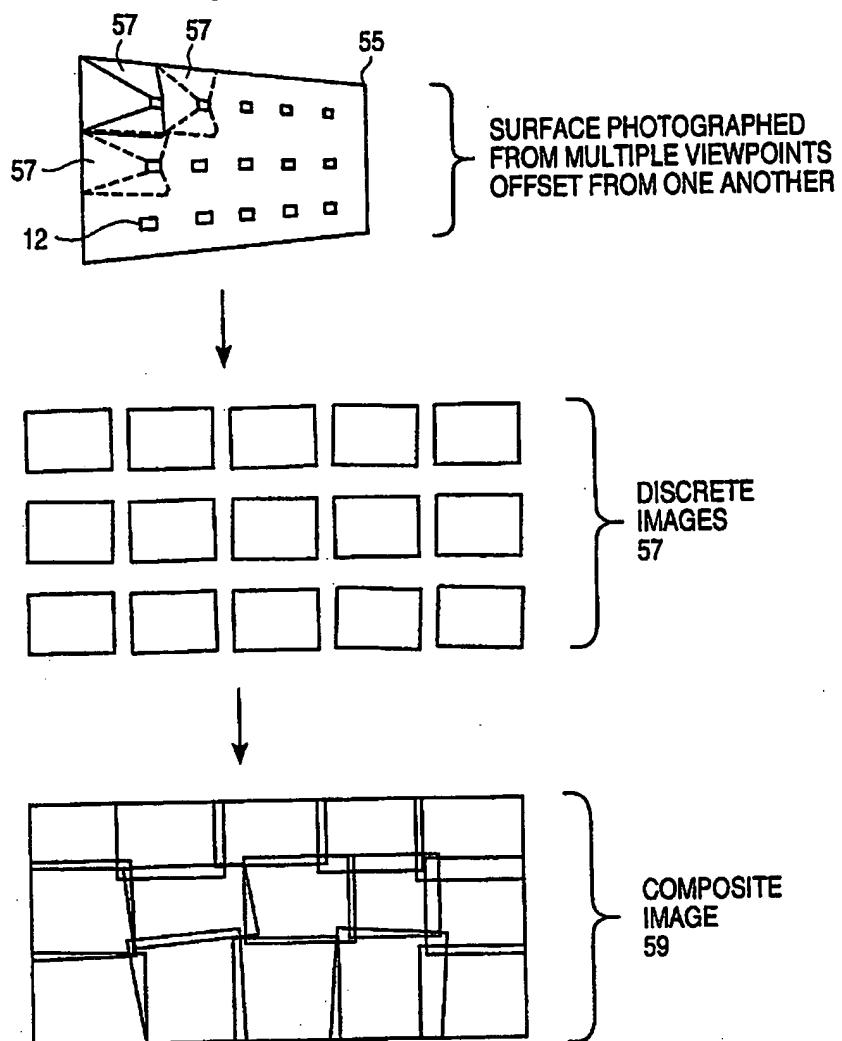


FIG. 3

SUBSTITUTE SHEET (RULE 26)

4/9

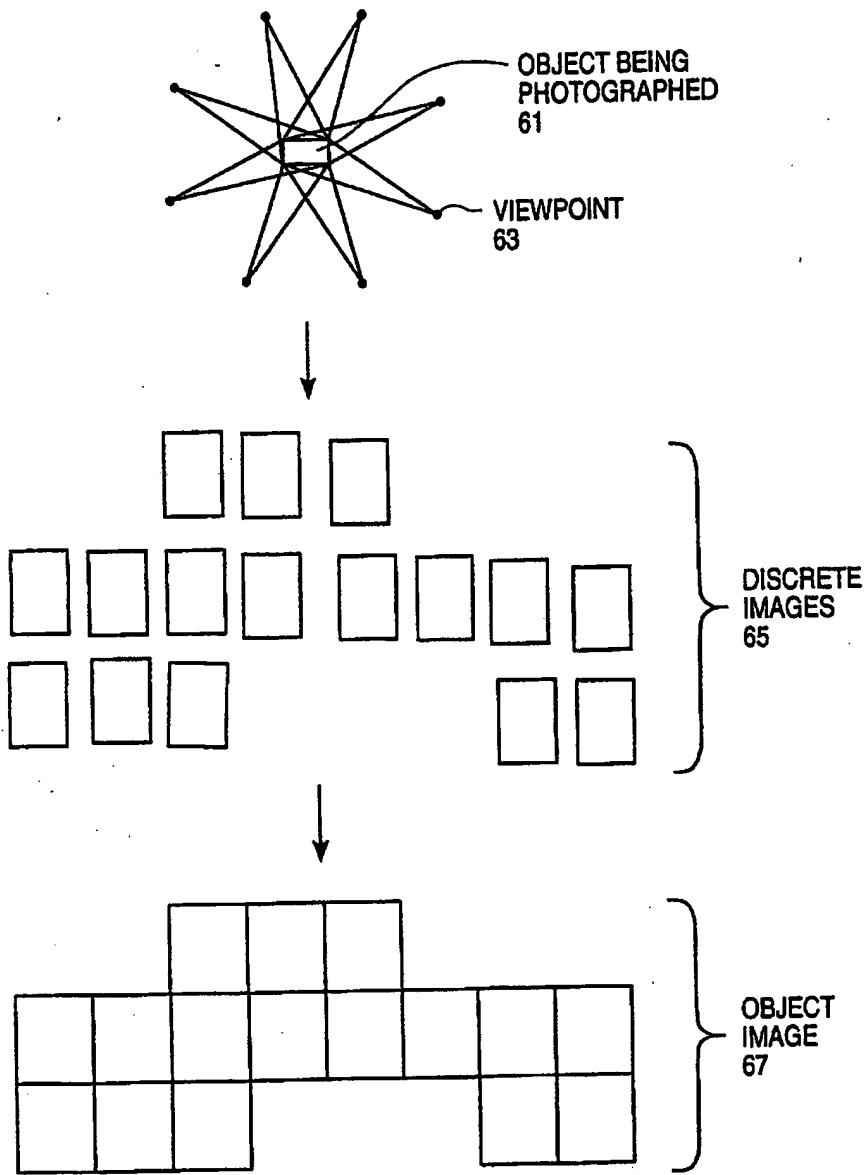


FIG. 4

SUBSTITUTE SHEET (RULE 26)

5/9

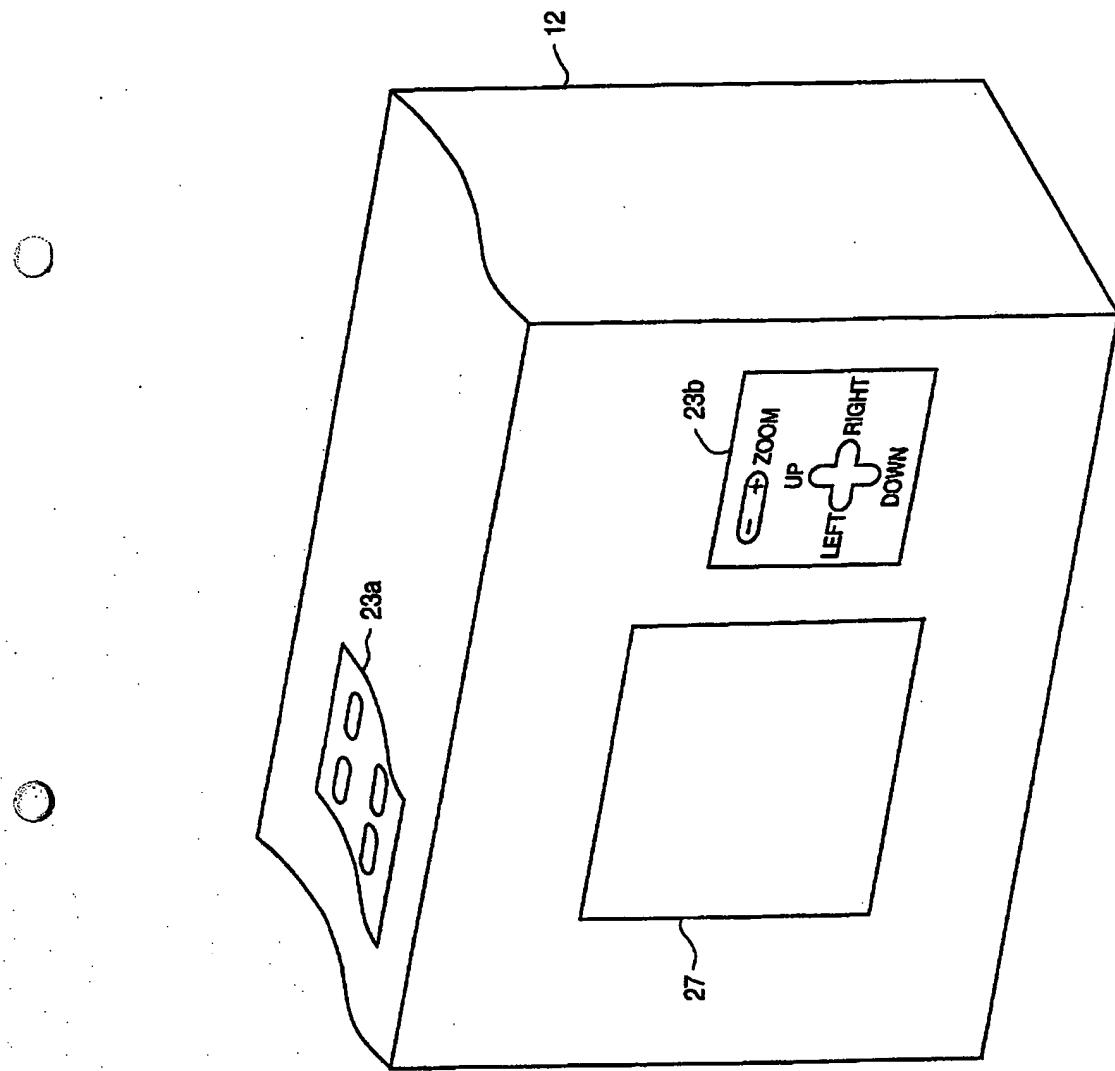


FIG. 5

SUBSTITUTE SHEET (RULE 26)

6/9

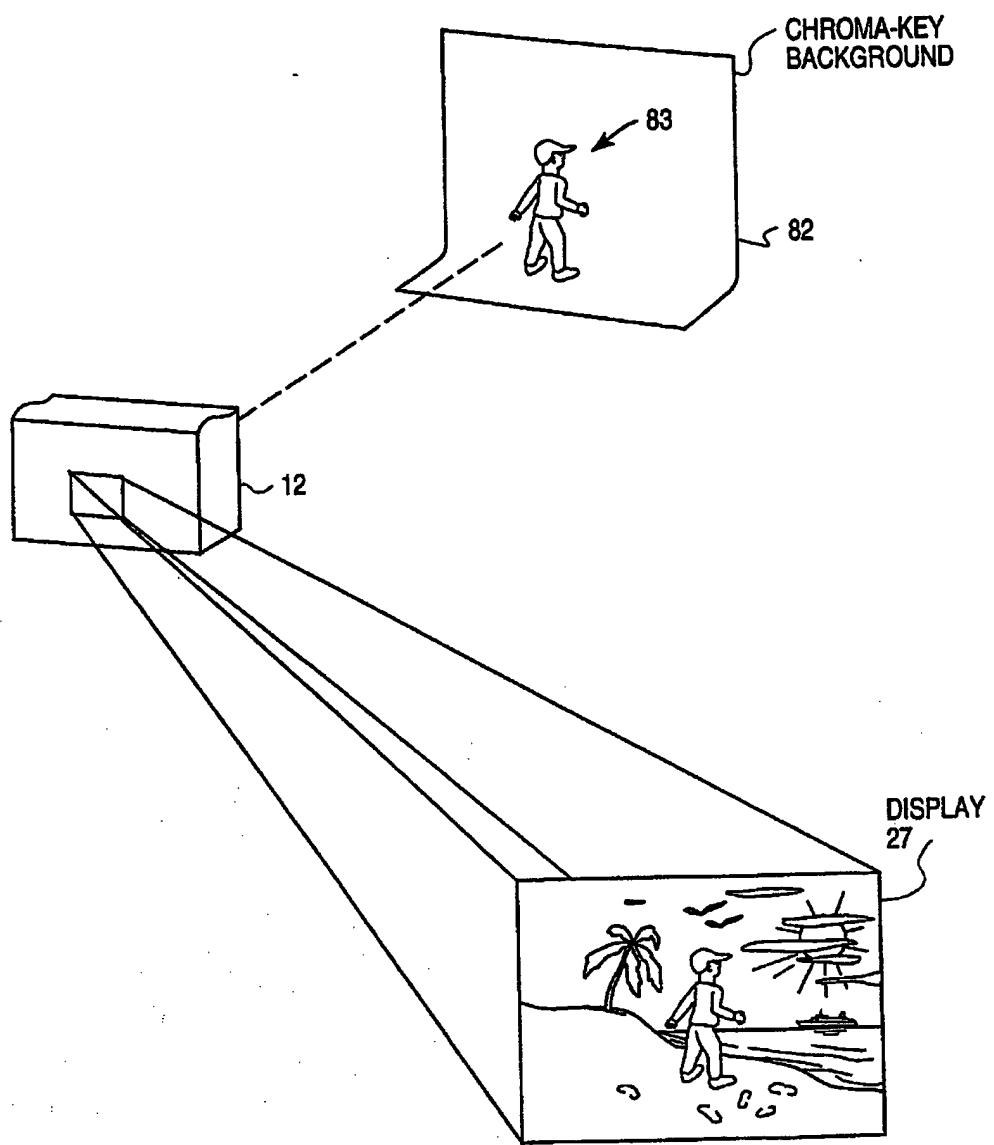


FIG. 6

SUBSTITUTE SHEET (RULE 26)

7/9

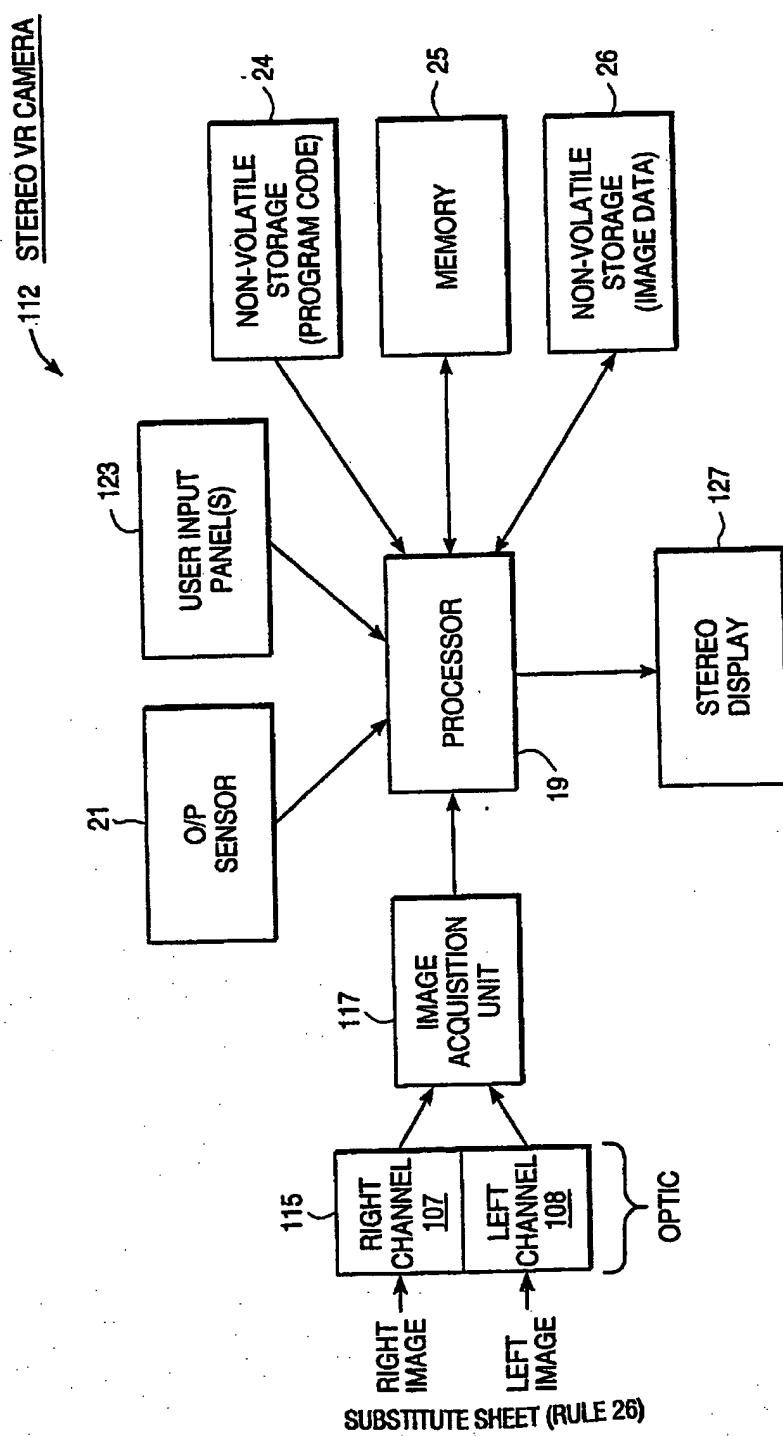


FIG. 7

8/9

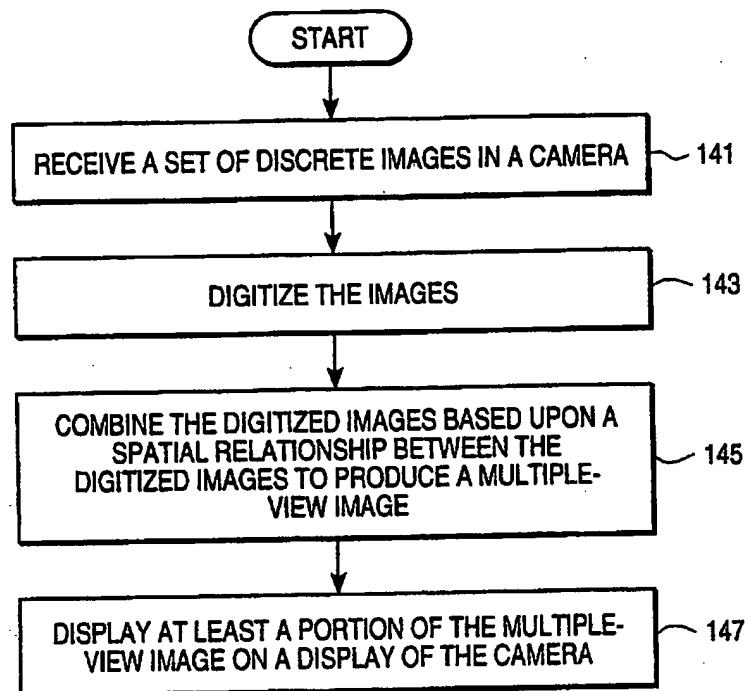


FIG. 8

9/9

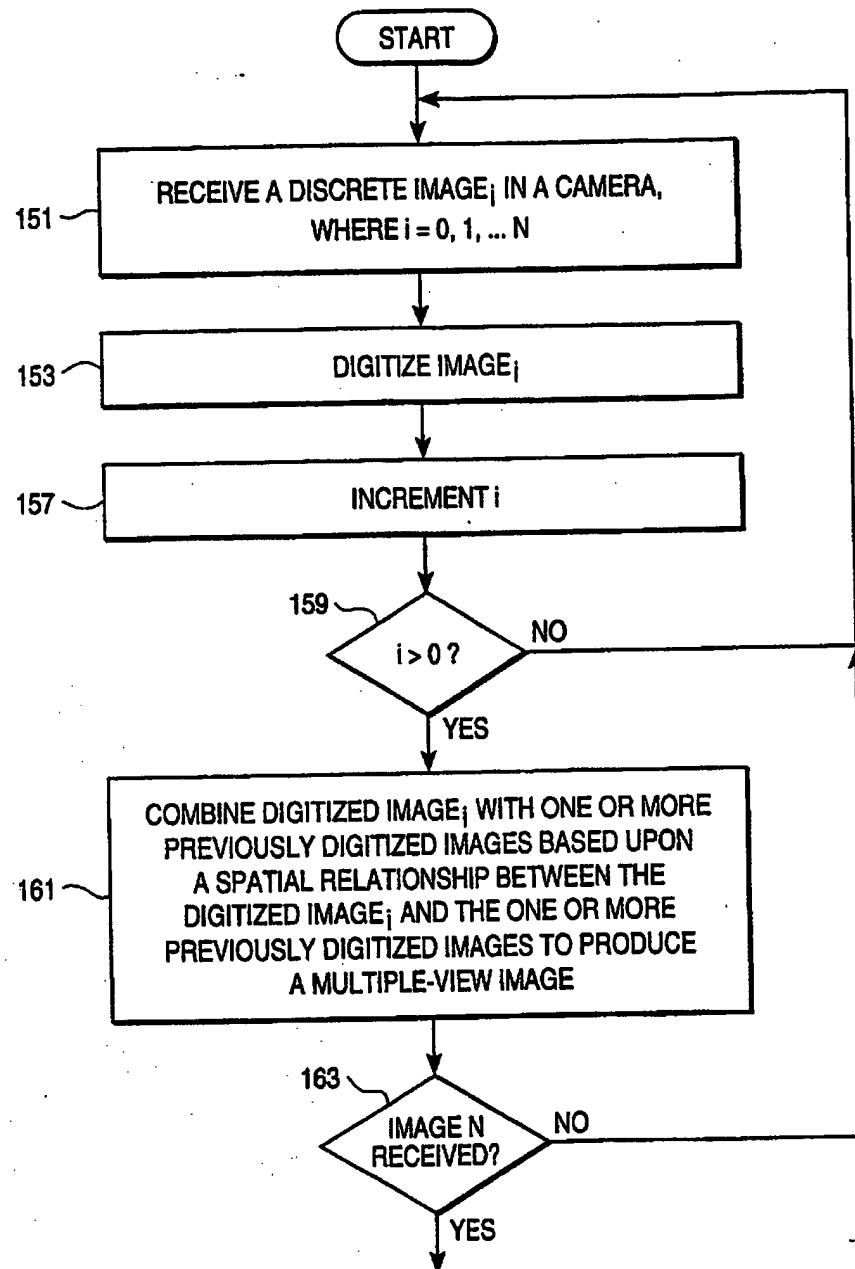


FIG. 9  
SUBSTITUTE SHEET (RULE 26)

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US98/13465

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :H04N 5/225  
US CL :348/218, 39

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 348/218, 39

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  
NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS  
camera?, panoramic, sensor or detector, stereo

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|-----------|--|-----------------------|
| X         | US 5,262,867 A (Kojima) 16 November 1993, col. 3, lines 23-68 and col. 4.          | 1-13, 17-34, 37-40    |
| —         |  | 14-16, 35, 36         |
| Y         |  |                       |
| X         | US 5,424,773 A (Saito) 13 June 1995, col. 2, lines 45-68 and col. 3.               | 1-13, 17-34, 37-40    |
| —         |  | 14-16, 35, 36         |
| Y         |  |                       |

Further documents are listed in the continuation of Box C.  See patent family annex.

|  |     |  |
|--|-----|--|
| * Special categories of cited documents: | "T" | later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention  |
| *'A'                                     |     | document defining the general state of the art which is not considered to be of particular relevance   |
| *'E'                                     |     | earlier document published on or after the international filing date   |
| *'L'                                     |     | document which casts doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reasons (as specified)  |
| *'O'                                     |     | document referring to an oral disclosure, use, exhibition or other sources   |
| *'P'                                     |     | document published prior to the international filing date but later than the priority date claimed   |
| "X"                                      |     | document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone   |
| "Y"                                      |     | document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art |
| "Z"                                      |     | document member of the same patent family  |

|  |   |
|--|---|
| Date of the actual completion of the international search<br><br>11 SEPTEMBER 1998 | Date of mailing of the international search report<br><br>19 OCT 1998 |
|--|---|

|   |   |
|---|---|
| Name and mailing address of the ISA/US<br>Commissioner of Patents and Trademarks<br>Box PCT<br>Washington, D.C. 20231<br>Facsimile No. (703) 305-3230 | Authorized officer<br><br>TUAN HO<br>Telephone No. (703) 305-4943 |
|---|---|

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US98/13465

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|-----------|--|-----------------------|
| X         | US 5,650,814 A (Florent et al) 22 July 1997, col. 7, lines 24-68 and col. 8.       | 1-13,17-34, 37-40     |
| -         |  | _____                 |
| Y         | US 5,646,679 A (Yano et al) 08 July 1997, col. 6, lines 31-68 and col. 7.          | 14-16, 35, 36         |
|           |  | 14-16, 35 and 36      |

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:



**BLACK BORDERS**

- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**